Taxonomies for Development

by Olivier Serrat

Taxis

Organizations spend millions of dollars on management systems without commensurate investments in the categorization needed to organize the information they rest on. Taxonomy work is strategic work: it enables efficient and interoperable retrieval and sharing of data, information, and knowledge by building needs and natural workflows in intuitive structures.

Bible readers think that taxonomy[1] is the world's oldest profession. Whatever the case, the word is now synonymous with any hierarchical system of classification that orders domains of inquiry into groups and signifies natural relationships among these. (A taxonomic scheme is often depicted as a "tree" and individual taxonomic units as "branches" in the tree.) Almost anything can be classified according to some taxonomic scheme. Resulting catalogs provide conceptual frameworks for miscellaneous purposes including knowledge identification, creation, storage, sharing, and use, including related decision making.

In their simplest expression, taxonomies are but systems for naming and organizing things. Not surprisingly, early conceptions applied to the living planet: Aristotle's animal classification, the first comprehensive attempt to compartmentalize that, divided organisms into two groups—plants and animals, the latter into blood and bloodless and then according to how they moved. Barring minor improvements, his system held well into the 18th century, when it w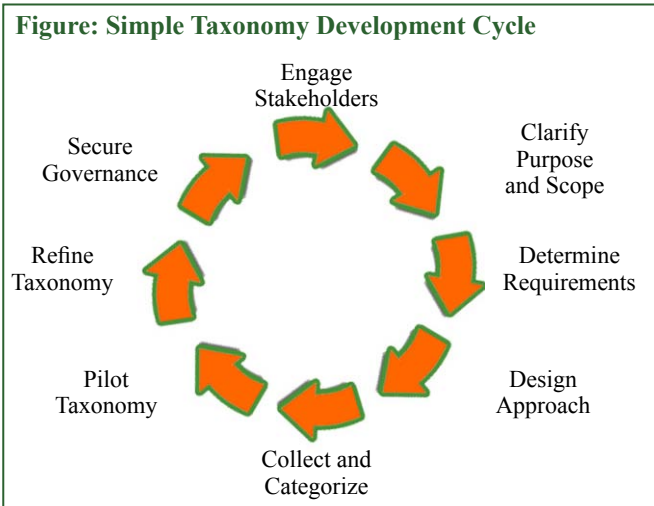as superseded by Linnaean taxonomy.[2] Librarians have coded and organized "books"—for example, in the form of blocks of wood, tablets, papyri, parchments, and papers—for a long time, too.[3]

> And out of the ground the LORD God formed every beast of the field, and every fowl of the air; and brought them to Adam to see what he would call them: and whatever Adam called every living creature, that was the name thereof.
>
> —Genesis 2:19

---

[1] The word "taxonomy" derives from the Greek *taxis* (signifying order or arrangement, from the verb *tassein*, meaning, to classify) and *nomos* (that is, law or science).

[2] Carl Linnaeus (1707–1778), a Swedish botanist, physician, and zoologist, laid the foundations for the modern scheme of nomenclature. He elaborated principles for defining genera and species of organisms and a uniform system for naming them: binomial (two-name) nomenclature—the first part is the genus, followed by the species. He is considered one of the fathers of modern ecology.

[3] Of course, almost anything—animate objects, concepts, events, inanimate objects, places, properties, and relationships—can be classified according to some taxonomic scheme. Some have explained that the human mind naturally organizes its knowledge of the world with such constructs, themselves shaped by local cultural and social systems.

The idea that information can be depicted as a tree is an old chestnut.[4] With the advent of the internet, however, the need to classify and categorize it has become even more urgent: beyond parent–child hierarchies, taxonomic schemes can now depict networks of relationships as well as the intensity of these.[5] Taxonomies will always matter because they help categorize information. In the digital age, however, information proliferates at a rate that far surpasses (traditional) institutional frameworks and controls. What is more, it can be classified with ease in myriad categories. (A book can only be placed in one place on a shelf.) Hence, organizations must move from nurturing trees to managing windswept piles of leaves under the watchful eye of the Semantic Web.[6] (Tags mean that users regulate information. Much as search engines do, this provides clues about what content has been deemed useful.) They must now work with taxonomies[7] every day to maximize the value and capability of their business, or function unsustainably under par for lack of ability to index, retrieve, organize, and help navigate knowledge assets. Taxonomy work is strategic work.

Figure: Simple Taxonomy Development Cycle

Engage Stakeholders — Clarify Purpose and Scope — Determine Requirements — Design Approach — Collect and Categorize — Pilot Taxonomy — Refine Taxonomy — Secure Governance

Source: Author

It's All About Context and Sense Making

Taxonomies are not artifacts for safekeeping: they adapt and change in coevolution with the efforts of users to make sense of ambiguity, emergence, and uncertainty in their environment.[8] (Therefore, effective taxonomies are extensible over time.) To begin, however, context drives information needs, which in turn spur the identification, creation, storage, sharing, and use of content. Logically, then, the availability of (or need for) content coupled with the information-seeking behaviors of users should influence the design and upkeep of taxonomies. Next, values or terms from the taxonomies must be applied to content.[9] (Logically, values or terms from the taxonomies should also be applied to staff profile pages.) And, as knowledge of information-seeking behaviors is garnered and staff profile pages expand, it becomes possible to both refine taxonomies and tag knowledge

---

4   The proverbial tree of knowledge bore many fruits: from hierarchies of organisms to offices and departments, we have fashioned life in tree-like ways, expediently forgetting that classification systems are neither value-free nor objectively true.

5   In Web 2.0 applications, tag (or word) clouds lead to collections of items that are associated with a particular tag. (Here, the notion of folksonomy is important: a folksonomy is a type of distributed system of categorization for ascribing and managing tags to online items such as images, videos, bookmarks, and text. Typically, users freely select tags from a chosen set of keywords, category names, or metadata. Examples of folksonomy systems are Delicious and Flickr, available at www.delicious.com and www.flickr.com, respectively.) Admittedly, low-quality tags are no replacement for formal systems: however, it may well be their emergent quality and openness that makes folksonomy tagging so useful; the tags that users select and attach can always be refined. There is also scope for improving tag literacy; the community would need to set rules and agree on standards for tags.

6   Rendering taxonomy web-based is inevitable and desirable in equal measure. The Semantic Web is a group of methods and technologies that would allow machines to understand the meaning, or "semantics", of information on the internet. Beyond Web 2.0's participatory technologies and social networks, the vision of the Semantic Web, aka Web 3.0, is to link metadata in such ways that it can easily be processed by machines on a global scale. It lies where computers, not humans, generate new information.

7   These *Knowledge Solutions* use the term "taxonomy" inclusively to refer to any classified collection of elements, be it descriptive—meaning labeled or tagged, or navigational—that is, aimed at facilitating the discovery of information through browsing. (Ontologies, which are used to reason about the properties of a domain and may be used to describe it, and mind maps, which involve considerable human interpretation, lie at the polar extremes of formality and potential for inference.)

8   That is a function of the organizational context, organizational knowledge, inter- and intra-organizational relationships, and the external environment. Taxonomy must be structured around people (both at the individual and at the community levels), their ideas and activities, and the information systems and technologies that are available to them.

9   In most instances, one can directly apply values or terms from taxonomies through metadata tags or by adding properties to files. One can also apply metadata tags indirectly by storing values or terms separately from content but by providing a pointer to that. (Content registries, metadata registries, and library catalogs are examples of indirect application.)

workers with metadata to identify them to other users and "push" what data, information, and knowledge will help them create situational awareness and understanding in complicated or complex situations so that they may make decisions.[10]

Table: Essential Steps to Taxonomy Design

Assign Roles (and Associated Responsibilities)
- The role of a governance board is to define strategy and appropriate types of content.
- The function of a taxonomy team, comprising 6–12 members, is to ensure the value of content placement and metadata.
- Content owners prepare content and apply metadata.
- Content managers edit and approve content.

Know Your Content
- Clean out obsolete content.
- Strive for topical taxonomy, with attention to scope, use, complexity, and scalability.
- Give every item one correct categorization.
- Accept that items can be organized in multiple categories.
- Minimize the number of "clicks".
- Build in flexibility and redundancy.
- Understand that it takes time to tag (or re-tag) content.

Before Getting Started, Understand Your
- Business context and priorities
- Knowledge workers
- Content
- Clients, audiences, and partners
- Information systems and technologies
- Limitations

Get Started
- Seek inspiration from existing taxonomic schemes: much taxonomy, unitary or not, already resides on the internet.
- Focus on primary, top-level concepts.
- Keep the taxonomy broad, simple, shallow, and elegant.
- Decide on standards values or terms that can be applied logically and consistently across different types of items. (Link synonyms and related terms.)
- Identify a general, intuitive category for the area of work being addressed.
- Define 6–12 top-level subcategories that are consistent with user expectations.
- Drill 2–3 levels deep.
- Repeat the process of division, based on the planned application of the taxonomy and the users concerned.
- Establish and share simple rules to encourage consistent practice and provide guidance on how to use different taxonomies.
- Review the draft taxonomy with users and subject matter experts.
- Test for user satisfaction from information-seeking tasks using the taxonomy.
- Refine and maintain the taxonomy, using it to the fullest.

Source: Adapted from Michael Pincher. 2010. A Guide to Developing Taxonomies for Effective Data Management. ComputerWeekly.Com. 8 March. Available: www.computerweekly.com/Articles/2010/04/06/240539/a-guide-to-developing-taxonomies-for-effective-data.htm

---

[10] Key applications, among many others, would include indexing, searching, and retrieval of project and program information and profiles of staff expertise on the internet, intranets, or shared drives.

![Knowledge Solutions logo]

Box: A Study of ADB's Knowledge Taxonomy

Effective use of data, information, and knowledge is essential to the development effectiveness of the Asian Development Bank.[a] Even so, how knowledge is identified, created, stored, and shared is as important as the use it is put to. A vital ingredient of that is a taxonomy with which to classify data, information, and knowledge in an ordered system that indicates natural relationships. In 2004, *Knowledge Management in ADB*[b] characterized the variety of knowledge products and services that ADB provides to its developing member countries and other stakeholders as (i) formal knowledge products and services, that are programmed as such and targeted at clients, audiences, and partners; and (ii) knowledge by-products derived from delivering loans or other activities.[c] To note, the framework document did not consider what data, information, and knowledge—including their flows—ADB rests on, that must be continuously enriched and facilitated to operate ADB itself as a learning organization.

More recent documents, expressly *Enhancing Knowledge Management under Strategy 2020: Plan of Action for 2009–2011*[d] and the attendant *Knowledge Management Results Framework*,[e] are testimony to the fact that ADB feels it can manage knowledge better, ultimately to the benefit of its developing member countries. However, in spite of noteworthy recent accomplishments, a sense of dissatisfaction remains due to the drive for unremitting improvement that is inherent to any knowledge strategy. One area of concern is ADB's knowledge taxonomy, which many if not all consider unrepresentative, unwieldy, and, frankly, uninformative.[f] With rapid changes in the demand for ADB's knowledge products and services in countries such as the People's Republic of China and India but elsewhere in general, ADB must think hard about how best to classify, describe, and map its most precious resource. (ADB's core assets might be described as financing, knowledge, and convening power.) Fast-transforming information systems and technologies, including the advent of Web. 2.0 and the Semantic Web, leave it no choice.

---

[a] Applications are well-nigh innumerable. They include policy and strategy formulation, strategic communication, business process formulation and implementation, corporate reporting, managing for results, staff learning and development, country partnership strategy formulation, policy dialogue, lending and nonlending activities, partnership building, effective knowledge management, project administration, and monitoring and evaluation for learning and accountability.

[b] ADB. 2004. *Knowledge Management in ADB*. Manila. Available: www.adb.org/knowledge-management/knowledge-framework.asp

[c] ADB's classification echoes a tendency to prioritize value-generating types of knowledge, specifically, the expertise that enables an organization to achieve its goals.

[d] Available: www.adb.org/knowledge-management/action-plan.asp. The action plan is reproduced in ADB. 2009. *Enhancing Knowledge Management Strategies*. Manila. Available: www.adb.org/documents/information/knowledge-solutions/enhancing-knowledge-management-strategies.pdf

[e] Available: www.adb.org/knowledge-management/results-framework.asp. The framework document is reproduced in ADB. 2010. *Crafting a Knowledge Management Results Framework*. Manila. Available: www.adb.org/documents/information/knowledge-solutions/crafting-a-knowledge-management-results-framework.pdf

[f] If taxonomies classify, describe, and map knowledge domains, taxonomy work is what one must do to achieve that outcome. The list of activities includes listing, creating and modifying categories, standardizing, mapping, representing, discovering native vocabularies and categories, and negotiating common terms. From this perspective, the classification proposed in *Knowledge Management in ADB* is unsophisticated. The application of taxonomies to organizations is more than the mere cataloging or indexing of documents. This is evident from the distinction between information management—which is the collection and management of information from one or more sources and the distribution of that information to one or more audiences—and knowledge management—which comprises a range of strategies and practices to identify, create, store, share, and use knowledge to support decisions and related tasks. From this, Patrick Lambe distinguishes three kinds of taxonomies: (i) objective taxonomies, usually of physical things, e.g., biological species, books; (ii) "embedded" taxonomies of how an organization has always done things (and often taken that for granted); and (iii) negotiated taxonomies based on stakeholder agreement or social negotiation. He sees that different kinds of taxonomies can (i) structure and organize (both things and processes), (ii) establish common ground, (iii) span boundaries between groups, (iv) help in sense making, and (v) aid in the discovery of risk and opportunity. See Patrick Lambe. 2007. *Organizing Knowledge: Taxonomies, Knowledge, and Organizational Effectiveness.* Neal-Schuman Publishers.

In 2010, the Knowledge Management Center in the Regional and Sustainable Development Department in ADB brainstormed on a possible classification of knowledge. It intuited that ADB's knowledge falls into four categories that marry tacit and explicit forms:

- **Lending and Nonlending Operations Knowledge.** This is largely tacit, even if official knowledge products are strictly codified and there is much potential to better learn before, during, and after implementation in more explicit ways.
- **Sector and Thematic Knowledge.** This is largely tacit, but communities and networks of practice increasingly offer ways to make that know-how explicit. More and more, strategic partnerships include knowledge components too.
- **Research Knowledge.** This is primarily explicit. However, staff also hold a wealth of tacit research know-how in their subject areas and research methods, as well as insights about how their work fits into the wider development context.
- **Business and Corporate Knowledge.** This is primarily explicit know-how about the corporate framework, for example, operational policies, operational procedures, project administration instructions, and business processes. Much codified project-management knowledge lies in databases. Tacit "street-wise" knowledge also exists.

The study of ADB's knowledge taxonomy purports to explore, recommend, and draw implications from a classification of knowledge products and services that improves ADB's organizational efficiency from internal and external perspectives. The classification contained in *Knowledge Management in ADB* and the possible taxonomy outlined by the Knowledge Management Center will be subjected to validation. The knowledge audit methodology described in *Auditing the Lessons Architecture*[g] may serve as reference.

- **Define the Scope, Purpose, and Types of Content Formats.** This entails also identifying the target audiences the taxonomy will serve, both internally and externally.[h] Sample questions to be addressed include: What objective does the proposed taxonomy hope to meet? What are the problems that staff in ADB are trying to solve and what concepts are important to them? What do they spend most of their time searching for? What are the existing sources for categorizing information? What are current, typical (or desired) information flows?[i] Are there technology constraints that might impact the development of the proposed taxonomy? What are the implications for ADB's external audiences? Accepting that knowledge is an asset while learning is a practice, what is the demand and supply for both tacit and explicit forms of knowledge, including the time dimensions of these?
- **Identify Concepts within the Proposed Taxonomy**. This entails discovering where and what the needed contents are, performing a content inventory, and conducting user and subject-matter expert interviews.[j]
- **Develop a Draft Taxonomy Organized Around Major Domains.** This entails establishing common rules for the format, relationships, and structure of taxonomy values or terms.
- **Review and Refine the Draft Taxonomy with Users and Subject-Matter Experts**. Sample questions to be addressed include: Are the users and subject-matter experts able to validate the taxonomy? Does the structure make sense to them? What are their thought processes and expectations? Are all major concepts included in the taxonomy? Does the taxonomy go too deep in any place? Are there any gaps?[k] Supported by a review of the literature and using best practices from other agencies as comparators, the

---

[g]   ADB. 2008. *Auditing the Knowledge Architecture*. Manila. Available: www.adb.org/documents/studies/auditing-lessons-architecture/in371-07.asp

[h]   A sample question to be addressed is: Can a picture of client demand be drawn, based on the inquiries sent by external "consumers" of ADB's knowledge, as expressed, for example, by hit and download trends through ADB.org?

[i]   Sample questions to be addressed include the following: What type of knowledge is needed? Who provides it and how does it arrive? How is it improved and reused? What happens to new knowledge that is created? What hinders ADB from doing more, better, faster? How can knowledge flows (therefore) be improved?

[j]   Since staff numbers and the range of subject matters in ADB are high, the approach to the conduct of interviews will involve a mix of representative interviews and the use of an online questionnaire.

[k]   Gaps in the suggested categories can be elicited through the online questionnaire.

approach to the study will focus on a series of interviews that will address the challenge from two angles:

- Using a typical decision as a starting point, tracking back to the factors that influence it to understand where ADB's knowledge products and services fit.
- Using ADB's knowledge products and services as starting points, tracking forward to understand the ways in which these are used and influence the work of ADB staff and other users of the said knowledge products and services.

Within the context of assessing the influence of knowledge on policy and practice, these methods tend to underestimate and overestimate, respectively, the state of affairs. Hence, by approaching the topic from both perspectives, it will be possible to work toward a balanced understanding based on triangulation of findings from interviewees (as well as between interviewees).

To cover the first angle, questions will draw on good practice for understanding real-world decision-making, as seen in works such as those of Cynthia Kurtz and David Snowden[l] and Gary Klein,[m] as well as a decision making typology recently developed for a study of how the United Kingdom's Department for International Development learns from research and evaluations. Questions will be fairly open to delve into how interviewees typically approach decisions in a real-world context, in order to draw out the ways in which they use different knowledge products and services in their everyday work. The interviews will be sensitive to a number of theoretical issues in this area that should aid the analysis, for example: different phases of policy decision making (e.g., agenda formulation, implementation, and evaluation); decision regimes (e.g., routine, incremental, fundamental, and emergent); and learning styles (e.g., activist, reflector, theorist, and pragmatist).

To cover the second angle, questions will be developed around a "theory of change" concerning how knowledge is used in ADB. This approach is suitable for understanding complex issues and involves identifying expectations about how ADB's knowledge products and services are expected to contribute to lesson learning and, eventually, better practice; through what processes and mechanisms; and with what intermediate outcomes. Analyzing whether these processes are functioning, or whether different intermediate outcomes are occurring, will help analyze the relevance and usefulness of the current knowledge taxonomy, and will provide crucial insights into how that could be improved—or how the current taxonomy might be better employed.

Where possible, the interviews will identify a few examples of ADB knowledge products and services, or particular policy processes, and delve into these in slightly more depth to provide examples in the final report. This approach will help deliver some components of a knowledge audit, viz., knowledge needs analysis and knowledge flow analysis—but not knowledge mapping nor a knowledge inventory. The knowledge inventory and mapping will need to be taken from the literature and provided by ADB. It is this inventory (and its knowledge taxonomy) that will be under review during the study.

---

[l]   Cynthia Kurtz and David Snowden. 2003. The New Dynamics of Strategy: Sense-Making in a Complex and Complicated World. *IBM Systems Journal*. Vol. 42, No. 3.
[m]   Gary Klein. 2001. *Sources of Power: How People Make Decisions*. MIT Press.

Source: Excerpted from ADB. 2010. *Terms of Reference for a Study of ADB's Knowledge Taxonomy*. Manila

Prologue

In 1998, Peter Drucker declared that the next information revolution would be in concepts, not technology, machinery, techniques, software, or speed.[11] It would ask: What is the meaning of information, and what is its purpose? That would lead rapidly to redefining the tasks to be done with the help of information and, with it, to redefining the institutions that perform these tasks. For learning organizations that leverage organization, people, knowledge, and technology, including the power of taxonomies, the revolution has already begun.

---

[11]   Peter Drucker. 1998. The Next Information Revolution. *Forbes*. 24 August. pp. 46–58.

Further Reading

ADB. 2008a. *Auditing the Lessons Architecture*. Manila. Available: www.adb.org/documents/studies/auditing-lessons-architecture/in371-07.asp

_____. 2008b. *Auditing Knowledge*. Manila. Available: www.adb.org/documents/information/knowledge-solutions/auditing-knowledge.pdf

_____. 2010. *Showcasing Knowledge*. Manila. Available: www.adb.org/documents/information/knowledge-solutions/showcasing-knowledge.pdf

Patrick Lambe. 2007. *Organizing Knowledge Taxonomies, Knowledge, and Organizational Effectiveness*. Chandos Publishing.

For further information

Asian Development Bank

ADB, based in Manila, is dedicated to reducing poverty in the Asia and Pacific region through inclusive economic growth, environmentally sustainable growth, and regional integration. Established in 1966, it is owned by 67 members—48 from the region. In 2007, it approved $10.1 billion of loans, $673 million of grant projects, and technical assistance amounting to $243 million.

Knowledge Solutions are handy, quick reference guides to tools, methods, and approaches that propel development forward and enhance its effects. They are offered as resources to ADB staff. They may also appeal to the development community and people having interest in knowledge and learning.